# Arnold's Canonical Matrices and the Asymptotic Simplification of Ordinary Differential Equations

Wolfgang Wasow*
*Department of Mathematics*
*University of Wisconsin—Madison*
*Madison, Wisconsin 53706*

Submitted by Hans Schneider

ABSTRACT

Let $A(x,\varepsilon)$ be an $n\times n$ matrix function holomorphic for $|x|\leqslant x_0$, $0<\varepsilon\leqslant\varepsilon_0$, and possessing, uniformly in $x$, an asymptotic expansion $A(x,\varepsilon)\sim\sum_{r=0}^{\infty}A_r(x)\varepsilon^r$, as $\varepsilon\to0^+$. An invertible, holomorphic matrix function $P(x,\varepsilon)$ with an asymptotic expansion $P(x,\varepsilon)\sim\sum_{r=0}^{\infty}P_r(x)\varepsilon^r$, as $\varepsilon\to0^+$, is constructed, such that the transformation $y=P(x,\varepsilon)z$ takes the differential equation $\varepsilon^h dy/dx=A(x,\varepsilon)y$, $h$ a positive integer, into $\varepsilon^h dz/dx=B(x,\varepsilon)z$, where $B(x,\varepsilon)$ is asymptotically equal, to all orders, to a matrix in a canonical form for holomorphic matrices due to V. I. Arnold.

## 1. INTRODUCTION

In [1] V. I. Arnold constructed what might be called a canonical form for holomorphic matrix valued functions under holomorphic similarity transformations. The purpose of this paper is to apply Arnold's result to the asymptotic theory of linear ordinary differential equations with a parameter of the form

$$\varepsilon^h\frac{dy}{dx}=A(x,\varepsilon)\,y, \tag{1.1}$$

where $A(x,\varepsilon)$ is an $n\times n$ matrix, holomorphic in both variables for

$$|x|\leqslant x_0,\qquad 0<\varepsilon\leqslant\varepsilon_0 \tag{1.2}$$

*Research supported by NSF grant MCS76-05979.

and having, uniformly for $|x| \leqslant x_0$, an asymptotic expansion

$$A(x,\varepsilon) \sim \sum_{r=0}^{\infty} A_r(x)\varepsilon^r \qquad \text{as} \quad \varepsilon \to 0+. \tag{1.3}$$

In Theorem 3.1 a kind of normal form for problems of the type (1.1) is described, which can be made the starting point of further analysis. The results are local, not global, which is also true of Arnold's theory.

## 2. ARNOLD'S CANONICAL FORM

This section is a brief summary of Arnold's main result in [1] specialized to one independent variable.

Let $M(x)$ be an $n \times n$ matrix holomorphic in $|x| \leqslant x_0$. By a constant similarity transformation, $M(0)$ can be changed into its Jordan form. We assume, without loss of generality, that this has been done already and that $M(0)$ is *upper* triangular.

Assumption I. *$M(0)$ is in Jordan canonical form.*

Thanks to an often proved theorem (see, e.g., [4]), $M(x)$ can be changed by a similarity transformation with a holomorphic matrix into a block diagonal matrix, holomorphic at $x=0$, each block of which has only *one* distinct eigenvalue at $x=0$.

Hence, the hypothesis below also does not entail a restriction of generality:

Assumption II. *$M(0)$ has only one distinct eigenvalue.*

Let $m_1 \geqslant m_2 \geqslant \cdots \geqslant m_p$ be the degrees of the elementary divisors of $M(0)$, and assume that the corresponding Jordan blocks are arranged in the same order of decreasing size. These blocks generate a partitioning of $M(0)$, the blocks of which may be called $M^{\mu\nu}(0)$, $\mu,\nu = 1,2,\ldots,p$. The same notation with superscripts will be used for the corresponding partitioning of any $n \times n$ matrix.

For a simple description of Arnold's canonical matrices we introduce a set of distinct $n \times n$ matrices $\Gamma_j$, $j = 1,2,\ldots,d$, defined as follows: Each $\Gamma_j$ has one entry of value 1, all other entries being zero. If $\Gamma_j^{\mu\nu}$ is the block (as described above) of $\Gamma_j$ which is not zero, then the nonzero entry is in the *last row* of the block if $\mu \leqslant \nu$, and in the *first column* if $\mu > \nu$. All such matrices are included in the set $\Gamma_1, \Gamma_2, \ldots, \Gamma_d$. For the sake of precision, we specify that the ordering of this set of matrices is to be such that $\Gamma_1$ has its nonzero entry in the first position of the last row of $\Gamma_1^{11}$ and that the numbering then proceeds from left to right and downward as in ordinary English writing. The number $d$ is given by the formula

$$d = \sum_{\mu=1}^{p} (2\mu - 1) m_\mu, \tag{2.1}$$

as can be readily verified.

This being clearly a case where a picture is worth a thousand words, Fig. 1 may be helpful.

Arnold's theorem, specialized to functions of one variable, says the following.

THEOREM 2.1 (Arnold). *Corresponding to every $n \times n$ matrix function $M(x)$, holomorphic at $x=0$ and satisfying Assumptions* I *and* II, *there exist $d$ scalar functions $\rho_j(x)$, $j = 1, 2, \ldots, d$, holomorphic and equal to zero at $x=0$, such that $M(x)$ is holomorphically similar to*

$$M(0) + \sum_{j=1}^{d} \rho_j(x) \Gamma_j \tag{2.2}$$

*in some neighborhood of $x=0$.*

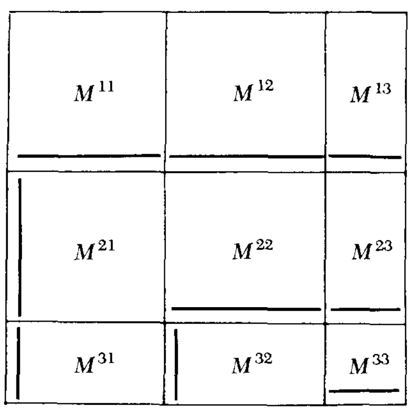


FIG. 1.

"Holomorphically similar" means "pointwise similar in a region by means of a transformation with a holomorphic matrix with a holomorphic inverse".

REMARK. The functions $\rho_j(x)$ are not always uniquely determined, but it is proved in [1] that no family of matrices involving fewer than $d$ independent parameter functions can replace (2.3) in the statement of Theorem 2.1. This fact justifies the name "Arnold's canonical matrices" for matrices of the form (2.2) or, more generally, for matrices that are direct sums of such matrices.

If $M(0)$ has many distinct eigenvalues, or if to a multiple eigenvalue there belong relatively few elementary divisors, Arnold's canonical form for $M(x)$ may be substantially simpler than $M(x)$ itself, in that it is then a rather sparse matrix (i.e., one having many zero entries). On the other hand, there are cases in which reduction to Arnold's form may hardly be worth while. In the extreme case that $M(0)$ is a multiple of the identity matrix, Arnold's theorem is vacuous: $M(x)$ itself is in canonical form.

For later use in this paper, I state one essential ingredient of Arnold's proof of Theorem 2.1 as a lemma.

LEMMA 2.1. *Every element $F$ in the vector space $\mathbf{C}^{n^2}$ of all $n\times n$ matrices with constant complex valued entries can be written (not uniquely) in the form*

$$F=PM(0)-M(0)P-\sum_{j=1}^{d}\rho_j\Gamma_j,$$

*where the matrix $P$ of $\mathbf{C}^{n^2}$ and the complex scalars $\rho_j$ depend on $F$.*

## 3. FORMAL SIMPLIFICATION OF THE DIFFERENTIAL EQUATION

Let $P(x,\varepsilon)$ be an $n\times n$ matrix function with the same properties as $A(x,\varepsilon)$ in Sec. 1; i.e., $P$ is holomorphic in (1.2) and admits a uniform asymptotic expansion

$$P(x,\varepsilon)\sim\sum_{r=0}^{\infty}P_r(x)\varepsilon^r,\qquad \varepsilon\to0+. \tag{3.1}$$

If $P$ is, in addition, nonsingular, the transformation

$$z=P(x,\varepsilon)\,y \tag{3.2}$$

takes the differential equation (1.1) into another one,

$$\varepsilon^h \frac{dz}{dx} = B(x,\varepsilon)z, \tag{3.3}$$

of the same type. Arnold's result will be used to construct a transformation (3.2) for which $B(x,\varepsilon)$ is very nearly in Arnold's form and, in particular, as sparse as possible.

By virtue of a theorem of Sibuya [2, 4], it is no loss of generality to adopt the hypothesis below.

Assumption III. *$A_0(0)$ is in Jordan form and has only one distinct eigenvalue.*

If $\lambda$ is the eigenvalue of $A_0(0)$, the simple transformation

$$\tilde{y} = \exp\{-\varepsilon^{-h}\lambda x\}\, y,$$

which is not of the form (3.2), produces a problem in which Assumption III is still true and the eigenvalue $\lambda$ is *zero*. For convenience this hypothesis will also be added; then we have

Assumption IV. *$A_0(0)$ is nilpotent.*

Furthermore, an application of Theorem 2.1 shows that a holomorphic change of the dependent variable $y$, independent of $\varepsilon$, will transform $A_0(x)$ into Arnold's canonical form, so that we may adopt, from the outset,

Assumption V. *$A_0(x)$ is in Arnold's canonical form.*

The matrices $A$, $B$ and $P$ from (1.1), (3.3) and (3.2) are related by the differential equation

$$\varepsilon^h \frac{dP}{dx} = PA - BP. \tag{3.4}$$

Conversely, any nonsingular matrix $P$ that solves (3.4) defines a transformation (3.2) from (1.1) into (3.3). In this section, (3.4) will be solved in the formal sense by a series in powers of $\varepsilon$, when $B(x,\varepsilon)$ is chosen as a formal series that is as much like an Arnold canonical matrix as possible. More precisely: In (3.4), replace $A$ by the series in (1.3), $P$ by the series in (3.1) and $B$ by

$$\sum_{r=0}^{\infty} B_r(x)\varepsilon^r, \tag{3.5}$$

where

$$B_0(x)=A_0(x)=A_0(0)+\sum_{j=1}^{d}\rho_{j0}(x)\Gamma_j,\qquad \rho_{j0}(0)=0; \tag{3.6}$$

$$B_r(x)=\sum_{j=1}^{d}\rho_{jr}(x)\Gamma_j,\qquad r>0, \tag{3.7}$$

with the $\rho_{jr}(x)$ to be determined as holomorphic scalar functions. The usual formal operations on the series are to produce, after collecting like powers of $\varepsilon$, the same series in powers of $\varepsilon$ in both members of (3.4). We satisfy the first of the resulting recursive sequence of equations, i.e.,

$$P_0(x)A_0(x)-A_0(x)P_0(x)=0,$$

by setting

$$P_0(x)=I. \tag{3.8}$$

The other conditions can then be written in the form

$$P_r(x)A_0(x)-A_0(x)P_r(x)-\sum_{j=1}^{d}\rho_{jr}(x)\Gamma_j=F_r(x),\qquad r\geqslant 1, \tag{3.9}$$

where $F_r(x)$ is a known holomorphic function if $P_s(x)$, $\rho_{js}(x)$ have already been determined as holomorphic functions for $s\leqslant r-1$, $1\leqslant j\leqslant d$.

Formula (3.9) is a set of $n^2$ scalar linear equations for $n^2+d$ scalar functions of $x$, namely the $n^2$ entries of $P_r(x)$ and the $d$ functions $\rho_{jr}(x)$, $j=1,2,\ldots,d$. From Lemma 2.1 we know that this set of equations has a solution at $x=0$, no matter what $F_r(0)$ might be. In other words, the linear operator from $\mathbf{C}^{n^2+d}$ into $\mathbf{C}^{n^2}$ defined by the left member of (3.9) has maximal rank $n^2$ at $x=0$. It follows from a version of the implicit function theorem (or else from elementary facts of matrix theory) that (3.9) has solutions $P_r(x)$, $\rho_{jr}(x)$, $j=1,2,\ldots,d$, which are holomorphic in a disk $|x|\leqslant x_2$, independent of $r$. Neither $P_r(x)$ nor the $\rho_{jr}(x)$ are, in general, unique.

This completes the proof of the theorem below.

THEOREM 3.1. *If Assumptions* III, IV *and* V *(which can be satisfied without losing generality) are true, there exist matrices* $P_r(x)$, $r=0,1,\ldots$ *with* $P_0(x)=I$ *and scalars* $\rho_{jr}(x)$, $j=1,2,\ldots,d$, $r=0,1,\ldots,$ *all holomorphic in a disk*

$|x| \leqslant x_2$, *such that the transformation*

$$z = \left( \sum_{r=0}^{\infty} P_r(x)\varepsilon^r \right) y \tag{3.10}$$

*takes*

$$\varepsilon^h \frac{dy}{dx} = \left( \sum_{r=0}^{\infty} A_r(x)\varepsilon^r \right) y \tag{3.11}$$

*formally into*

$$\varepsilon^h \frac{dz}{dx} = \left[ A_0(0) + \sum_{j=1}^{d} \left( \sum_{r=0}^{\infty} \rho_{jr}(x)\varepsilon^r \right) \Gamma_j \right] z. \tag{3.12}$$

As was pointed out before, the functions $\rho_{jr}(x)$, are in general not unique. My result in [3] can be interpreted, in the present terminology, as a proof that all $\rho_{jr}(x)$, $r > 0$, may be chosen as zero in the special case when $p = 1$ (and therefore, $m_1 = n$, $d = n$), and also $\rho_{10}(0) \neq 0$.

## 4. REMARKS ON THE ANALYTIC SIMPLIFICATION

The analytic implications of the formal theorem 3.1 are not immediately apparent. It is true that there always exist matrix functions $P(x,\varepsilon)$ and scalar functions $\rho_j(x,\varepsilon)$ which are asymptotically represented by the formal series in (3.10) and (3.12) as $\varepsilon \to 0+$. This follows from the so-called Borel-Ritt theorem (see, e.g., [4], Theorem 9.6). As can be immediately verified, the transformation

$$z = P(x,\varepsilon)\, y$$

then changes (1.1) into

$$\varepsilon^h \frac{dz}{dx} = \left[ A_0(0) + \sum_{j=1}^{d} \rho_j(x,\varepsilon)\Gamma_j + E(x,\varepsilon) \right] z \tag{4.1}$$

with a matrix $E(x,\varepsilon)$ whose asymptotic expansion is zero, as $\varepsilon \to 0+$, uniformly for $|x| \leqslant x_2$. Nothing else is known about $E(x,\varepsilon)$.

Equation (4.1) is of some interest, but it falls short of what one wishes to achieve, unless it can be shown that the term $E(x,\varepsilon)$ causes only asymptotically negligible changes in the solutions of the differential equation. This task turns out to be equivalent to proving that $E(x,\varepsilon)$ can be made identically zero by a judicious choice of the functions $P(x,\varepsilon)$, $\rho_j(x,\varepsilon)$ among the infinitely many functions with the same asymptotic expansion.

The most general result in this direction known to me is contained in [5], [6]. These papers discuss in some generality the question as to when the existence of a *formal* transformation of (1.1) implies the existence of a corresponding *analytic* transformation. The sufficient condition for this to be the case, given in [5] (Hypothesis (H) of [5]) is, indeed, satisfied when $p=1$, $\rho_{10}(0)\neq 0$, the special case mentioned above. It is easy to describe numerous additional special types of differential equations for which [5], [6] establishes the analytic validity of our formal procedure.

The hypothesis (H) of [5] is, however, unnecessarily restrictive. A generalization of [5], [6] which applies to the present problem under milder hypotheses will be published separately.

## REFERENCES

1 V. I. Arnold, On matrices depending on parameters, *Usp. Mat. Nauk* **26**, No. 2 (1971), 101–114. Also English transl. in Russ. *Math. Surv.* **26**, No. 2 (1971), 29–43.
2 Y. Sibuya, Sur réduction analytique d'un système d'équations différentielles ordinaires linéaires contenant un paramètre, *J. Fac. Sci. Univ. Tokyo* **7**, No. 1 (1958), 527–540.
3 W. Wasow, Simplification of turning point problems for systems of linear differential equations. *Trans. Am. Math. Soc.* **106** (1963), 100–114.
4 W. Wasow, *Asymptotic Expansions for Ordinary Differential Equations*, Krieger, Huntington, N.Y., 1976 (orig. Wiley, 1965).
5 W. Wasow, On the analytic validity of formal simplifications of linear differential equations, part I, *Funkc. Ekvacioj* **9** (1966), 83–91.
6 W. Wasow, —, part II, *Funkc. Ekvacioj* **10** (1967), 107–122.

*Received 8 August 1976; revised 18 August 1976*